Subgradient methods with space transformation

Stetsyuk P.I., Zhydkov V.O.
stetsyukp@gmail.com

Department of non-smooth optimization methods
V.M. Glushkov Institute of Cybernetics, Kyiv, Ukraine

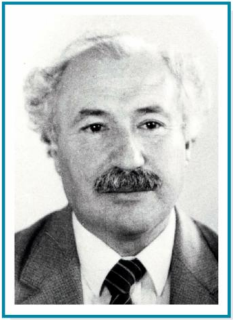
V.M. Glushkov Institute of Cybernetics, Kyiv, Ukraine, August 04, 2020

Outline

1. Shor's non-smooth optimization methods
2. Subgradient methods with space transformation
3. Packing 3D-objects (placement problem)
4. ED- and UC- energy problems
5. Search for defects in regular structures

Outline

Naum Z. Shor (1937–2006)



Shor founded scientific school of non-smooth optimization methods (Institute of Cybernetics, Kyiv, Ukraine)

In 1962 he developed the first subgradient method

In 1969 he used the operator of space dilation for acceleration of gradient methods convergences

N.Z. Shor's methods

are still of a great theoretical and applied importance, and "a key" for solution of large-scale problems.

Prof. Boyd's letter ‣stanford.edu/ boyd/

April 15, 2005

Dear Professor Shor,

We have never met, but your work has very much influenced me for many years now. I started with your small 1985 Springer book on subgradient methods, which I read as a PhD student. I recently read your newer book on nondifferentiable optimization (1998), which I enjoyed very much.

I'm enclosing copies of the three books I've written. The first concerns the design of linear controllers via convex optimization; the second is on linear matrix inequalities; and the third one is a basic textbook on convex optimization. [...] I hope you can see your strong influence in all of these books.

With the best regards,
Stephen P. Boyd

Shor's basic monographies

1. Шор Н.З. *Методы минимизации недифференцируемых функций и их приложения.* Киев: Наукова думка, 1979.

English translation: Shor N.Z. *Minimization Methods for Non-Differentiable Functions.* Berlin: Springer-Verlag, 1985.

2. Shor N.Z. *Nondifferentiable optimization and polynomial problems.* Boston; Dordrecht; London: Kluwer Academic Publishers, 1998.

3. Шор Н.З., Стеценко С.И. *Квадратичные экстремальные задачи и недифференцируемая оптимизация.* Киев: Наукова думка, 1989.

Outline

Subgradient methods with space transformation

purposed to solve the following problem

$$f^* = f(x^*) = \min_{x \in R^n} f(x),$$

where $f(x)$ – a convex function (smooth, non-smooth)

We have: x_0 – starting point, B_0 – $n \times n$-matrix,

Iterations $k=1, 2, \ldots$ have the following form

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}, \quad B_{k+1} = B_k T_k, \qquad \text{(Shor69)}$$

where h_k – step-size, $g_f(x_k)$ – a subgradient of function $f(x)$ at the point x_k, T_k – $n \times n$-matrix.

The most known methods

1. r-**algorithms** (robust to accumulation of errors);
2. **ellipsoid methods** (convergence - geometric progression);
3. **subgradient methods with space transformation** (use Fejer-type steps, Polyak's steps and others).

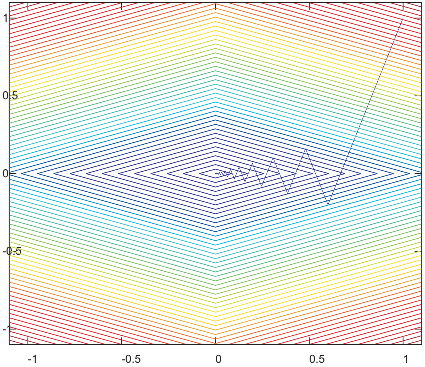
The methods have accelerated convergence

for convex ravine functions (smooth and non-smooth).

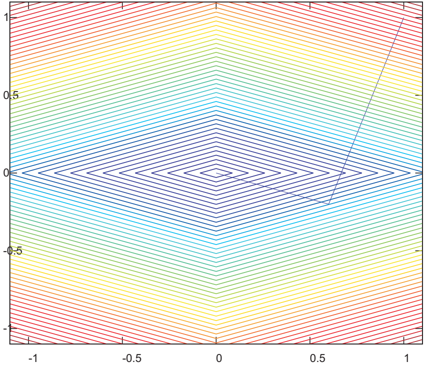
Implementations of the methods in Octave:

1. **ralgb5a** (30 rows), 2. **emshor** (15), 3. **amsg2p** (25).

Accelerated convergence: piecewise linear function



1. Trajectory of method **ams**



2. ...accelerated method **amsg2**

for ravine function $f_1(x_1, x_2) = |x_1| + 10|x_2|, \; x_0 = (1, 1).$

Implementations of r-algorithms by

N.Zhurbenko, L.Shabashova, A.Kuntsevich, A.Lykhovyd

were used for solving:

1. large-scale block optimization problems with various decomposition schemes;
2. minimax and matrix optimization problems;
3. for calculating dual bounds in multiextremal and combinatorial optimization problems;
4. ... and others

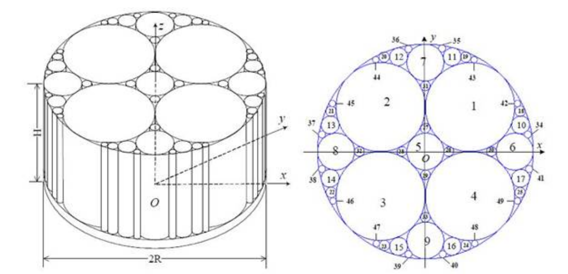
They were core for application packages (Kiev, CaSA)

PLANER (1983, **19**, 362–382), DISNEL (1991, **27**, 354–366).

Outline

Optimal balanced packing

project NASU–STCU (2012–2014)



1. Stetsyuk P., Romanova T., Scheithauer G. *On the global minimum in a balanced circular packing problem* // Optimization Letters. – 2016, № 10. – P. 1347–1360.

Formulation of the problem

Let circles S_i with radii r_i and weights w_i, $i = 1, \ldots, m$ be given. The center of mass for the circle S_i matches its center.

A balanced packing of a family of circles S_i, $i = 1, \ldots, m$ is called such a packing in the circle S, so that the radius of the circle S is minimal and the center of mass of the family of circles S_i, $i = 1, \ldots, m$, coincided with the center of S.

Need

to find balanced packing of the family of circles S_i, $i = 1, \ldots, m$.

Formulation of the problem

Multiextremal nonlinear programming problem:

$$r^* = \min_{r \geq 0} \ r \tag{1A}$$

subject to

$$x_i^2 + y_i^2 \leq (r - r_i)^2, \quad i = 1, \dots, m, \tag{2A}$$

$$(x_i - x_j)^2 + (y_i - y_j)^2 \geq (r_i + r_j)^2, \quad 1 \leq i < j \leq m, \tag{3A}$$

$$\sum_{i=1}^{m} w_i x_i = 0, \quad \sum_{i=1}^{m} w_i y_i = 0. \tag{4A}$$

Unknowns: x_i, y_i – center of the circle S_i, r – radius of the circle S; knowns: radius r_i and weight w_i of the circle S_i, $i = 1, \dots, m$.

Outline

Problems of power units loading

for Heat Power Plants (2012–2018)

2. Стецюк П.І., Журбенко М.Г., Лиховид О.П. *Математичні моделі та програмне забезпечення в задачах енергетики.* – К.: ПП Ательє "Поліграфічний комплекс", 2012.

„Optimal" loading for power units

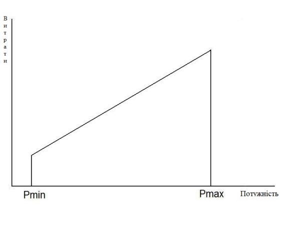
1. Economic Load Dispatch Problem (ED-problem)

We have N – a set of running power units. It is necessary to find electric load for each power unit (the amount of electricity generated by the power unit) such as to meet the needs of consumers E_t, $t=1,\ldots,T$ with the minimum total cost of electricity generation.

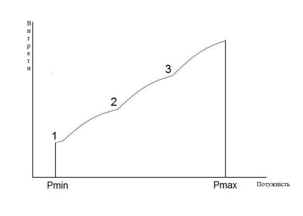
2. Unit Commitment Problem (UC-problem)

We have N_{UC} – a set of power units, which can be turned on/off. It is necessary to find intervals $t \in \{1,\ldots,T\}$, when each power unit should be turned on/off, and such electric load for each committed power unit that meets the needs of consumers E_t, $t=1,\ldots,T$ with the minimum total cost of electricity generation.

Functions of fuel consumption



linear (convex)



non-convex

Piece-concave function (with „valve“ effect):

$$f_i(x_{i,t}) = a_i x_{i,t}^2 + b_i x_{i,t} + c_i + \left| d_i \times \sin(\omega_i \times (P_i^{low} - x_{i,t})) \right|$$

Multiextremal ED-problem

$$\min \sum_{t=1}^{T} \sum_{i \in N} \left(a_i x_{i,t}^2 + b_i x_{i,t} + c_i + \left| d_i \sin(\omega_i (P_i^{low} - x_{i,t})) \right| \right) \quad (1B)$$

$$\sum_{i \in N} x_{i,t} = E_t, \quad t = 1, \dots, T, \quad (2B)$$

$$P_i^{low} \le x_{i,t} \le P_i^{up}, \ i \in N, \ t = 1, \dots, T, \quad (3B)$$

$$x_{i,t} - x_{i,t-1} \le U_i, \ i \in N_R \subset N, \ t = 2, \dots, T, \quad (4B)$$

$$x_{i,t-1} - x_{i,t} \le D_i, \ i \in N_R \subset N, \ t = 2, \dots, T. \quad (5B)$$

N_R – power units with constraints on maximal load change for time intervals $t-1$ and t („ramp-rate limits constrains“)

Mixed ED- and UC- problems (MIP-problem)

$$\min \sum_{t=1}^{T} \sum_{i \in N \setminus N_{UC}} (b_i x_{i,t} + c_i) + \sum_{t=1}^{T} \sum_{i \in N_{UC}} (b_i x_{i,t} + c_i y_{i,t}) \quad (1C)$$

$$\sum_{i \in N} x_{i,t} = E_t, \quad t = 1, \dots, T, \quad (2C)$$

$$P_i^{low} \le x_{i,t} \le P_i^{up}, \ i \in N \setminus N_{UC}, \ t = 1, \dots, T, \quad (3C)$$

$$P_i^{low} y_{i,t} \le x_{i,t} \le P_i^{up} y_{i,t}, \ i \in N_{UC}, \ t = 1, \dots, T, \quad (4C)$$

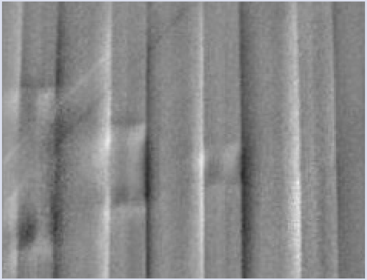
$$y_{i,t} \in \{0, 1\}, \ i \in N_{UC}, \ t = 1, \dots, T, \quad (5C)$$

$$x_{i,t} - x_{i,t-1} \le U_i, \ i \in N_R \subset N \setminus N_{UC}, \ t = 2, \dots, T, \quad (6C)$$

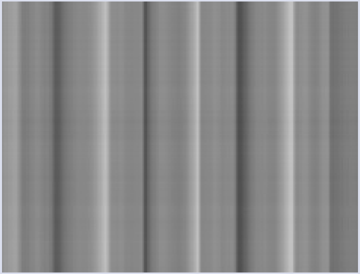
$$x_{i,t-1} - x_{i,t} \le D_i, \ i \in N_R \subset N \setminus N_{UC}, \ t = 2, \dots, T. \quad (7C)$$

Outline

What are defects in regular images



Regular, with defects



Regular, without defects

Defects

are defined as the difference between the left and right images

Regular Image Search Problem

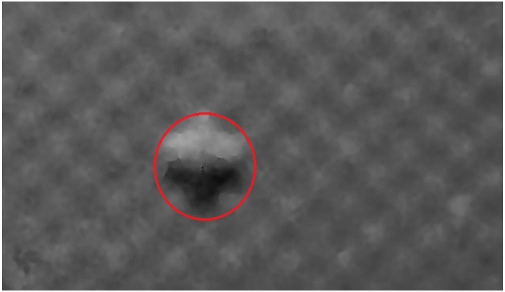
We have matrix $A \in R^{m \times n}$. It is necessary to find such vectors $x^* \in R^m$ та $y^* \in R^n$, that coefficients of matrix A^*: $a^*_{ij} = x^*_i + y^*_j$ minimally deviated from the coefficients of the matrix A, that is

$$(x^*, y^*) = \operatorname*{argmin}_{\substack{x \in R^m \\ y \in R^n}} \left\{ \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij} - x_i - y_j| + \left| \sum_{i=1}^{m} x_i - \sum_{j=1}^{n} y_j \right| \right\} \quad (1D)$$

the solution of the problem of minimizing a non-smooth function.

3. Стецюк П.И., Савицкий В.В. О поиске дефектов в регулярных 3D-структурах // Проблемы управления и информатики. – 2018. – № 2. – С. 33–48.

Regular rectangular images



sizes like 400×600 can be processed in dialog mode because the solution of the problem (1D) requires a few seconds on modern personal computers (GNU Octave v. 3.6.4).

*"A key" for solving large-scale problems

r-algorithms are successful in solving problems

QP ($n \gg m \approx 500,\ d \geq 0$) та LP ($m \gg n \approx 500$)

Quadratic Program (n,m)

$$\min_{x \in R^n} \sum_{i=1}^{n} (d_i x_i^2 + c_i x_i) \quad (1Q)$$

$$\sum_{i=1}^{n} a_{ij} x_i = b_j,\ \ j = \overline{1,m}, \quad (2Q)$$

$$x_i^{low} \leq x_i \leq x_i^{up},\ \ i = \overline{1,n}. \quad (3Q)$$

Linear Program (n,m)

$$\min_{x \in R^n} \sum_{i=1}^{n} c_i x_i \quad (1L)$$

$$\sum_{i=1}^{n} a_{ji} x_i \leq b_j,\ \ j = \overline{1,m}, \quad (2L)$$

$$x_i^{low} \leq x_i \leq x_i^{up},\ \ i = \overline{1,n}. \quad (3L)$$

Remark: to QP and LP, we can use numerical linear algebra

utilising Huawei's specific hardware for acceleration.

Questions?

THANK YOU FOR ATTENTION!

e-mail: *stetsyukp@gmail.com*